Increasing Access to the Results of Federally Funded Scientific Research

Office of Science and Technology Policy
Memorandum of February 22, 2013

Origin and scope

- Essentially a response to balance two poles of the vigorous "Open Access" debate:
  - 'Open Access Community'
    - Research scientists, institutional libraries, nonprofit publishers (e.g., PLOS), & others
    - 'We the People' petition
  - 'Traditional' for-profit scientific journal publishers
- Policy issued as a Memorandum rather than Executive Order
- Patterned on NIH data sharing policy
- Pending legislation: 'The Fair Access to Science and Technology Research (FASTR) Act of 2013'
- Applies to agencies with >$100m annual R&D expenditures

Goals of policy

- Ensure that "to the *greatest extent* and with the *fewest constraints possible* . . . the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community".

- Promote "wider availability of peer-reviewed publications and scientific data in digital formats" to "accelerate scientific breakthroughs and innovation, promote entrepreneurship, and enhance economic growth and job creation" and

- Create "innovative economic markets for services related to curation, preservation, analysis, and visualization"

- Note emphasis on economic goals and role of private sector

Included

- "Any results published in peer-reviewed scholarly publications" based on Federally-funded research, including:
  - *Publication* itself
  - *Supporting data*
    - Defined as "the digital recorded factual material commonly accepted in the scientific community as necessary to validate research findings *including data sets used to support scholarly publications*"
    - In *digital form* as opposed to embedded or 'entangled' in an inaccessible form (e.g., a data table in a PDF or paper report)
  - *Supplemental materials*, e.g., videos, images, diagrams, etc.

Excluded

- Classified research data & reports
- Laboratory (& field) notebooks
- Preliminary analyses
- Drafts of scientific papers still in review
- Plans for future research
- Peer review reports
- Communications with colleagues
- Physical objects, such as laboratory specimens

Agency public access plan

For *scientific publications* & *digital scientific data* created with Federal Funds, a strategy to:

- Leverage existing public & private-sector scholarly literature archives, e.g.,
  - PubMed, JSTOR, SciVerse/Scopus, USGS Pubs Warehouse
- Improve public discovery & access to digital data, e.g.,
  - General & thematic *portals*, *catalogs*, & *repositories*, e.g., Data.gov, Ocean.data.gov, ScienceBase, OBIS, Dryad, GBIF, IEDA
- Optimize search, archival, & dissemination approaches that encourage "innovation in accessibility and interoperability"
  - Web services, 'apps', linked data, semantic web
- Ensure long-term stewardship of data & information

Agency public access plan

Implementation & governance elements:

- Provide guidance to researchers on obligations
- Make procedural & policy changes as needed
- Put in place means to measure & enforce compliance
- Timeline for implementation
- Identify resources *within existing agency budgets*
- Identify special circumstances preventing implementation of any or all elements of the plan
- Solicit input from 'stakeholders'
  - Researchers, universities, libraries, publishers, users, etc.
- OSTP & OMB approval

Access to scientific publications

- "Any results published in *peer-reviewed scholarly publications*" based on Federally-funded research
- Applies to public release of:
  - *Final peer-reviewed manuscript* versions of research papers
  - *Supporting data*
  - *Associated content* (images, video, diagrams, etc.)
- *Twelve-month post-publication* embargo period
  - Treated as a *guideline*
  - Embargo period may be changed based on:
    - Agency's unique circumstances or mission needs;
    - Stakeholder demonstration of non-conformance of open access objectives

Access to scientific publications

Other requirements, hopes, & desires:

- Procedures to *prevent mass redistribution of scholarly publications* (e.g., the MIT/JSTOR incident)
- Provide *full, free digital access* post embargo
- Facilitate easy public search, analysis, & access (download)
- Ensure *full public access to metadata*
  - Metadata format (& standards) for "interoperability with current and future search technology"
- Metadata should provide link to post embargo location of full text & associated supplemental materials
- Access is Section 508 compliant

Access to scientific publications

Other requirements, hopes, & desires:

- Encourage *collaboration to maximize interoperability* between public & private search, access, & delivery platforms & repositories, e.g., ScienceDirect, JSTOR
- Ensure *attribution to author, journals, & original publishers* is maintained
- Ensure publications & metadata are *properly archived in repositories* (maintained by Federal agencies, consortia, scholarly or professional associations, publishers, libraries)
  - For *long-term preservation & access* to content *without charge*
  - Using widely available & where possible *nonproprietary archival formats*

Access to digital scientific data

Maximize *public access to digitally formatted scientific data* created with Federal funds *without charge* while:

- Protecting *personal privacy*
- Recognizing *proprietary interests, business confidential information, & intellectual property rights*
- Avoiding significant negative impact on *intellectual property rights*
- Balancing *value of long-term preservation & access and the associated cost & administrative burden*
- Providing for *appropriate attribution* (and credit)

Access to digital scientific data

Other elements:

- *Data management plans* required
- Inclusion of *appropriate costs for DM & access* in proposals
- Mechanisms to *assure compliance*
- *Partnerships* to improve data access & compatibility
- *Promote deposit of data associated with publications in publicly accessible databases*
  - Institutional and/or consortium repositories
- Support *training, education, & workforce development*
  - All aspects of the scientific data management lifecycle

Required actions & deadlines

- Draft plan due to OSTP by NLT *August 23, 2013*
- OSTP & OMB review & guidance on final plan
  - *No set date for final approval* of plan
- USGS Executives with reporting responsibilities:
  - Data: Kevin Gallagher (CSS)
  - Publications: Alan Thornhill (OSQI)

Loopholes and potholes

Barbossa [to Elizabeth]: First, your return to shore was not part of our negotiations nor our agreement so I must do nothing. And secondly, you must be a pirate for the pirate's code to apply and you're not. And thirdly, the code is more what you'd call "guidelines" than actual rules. Welcome aboard the *Black Pearl*, Miss Turner.   --Pirates of the Caribbean: The Curse of the Black Pearl (2002)

- *Unfunded mandate*: Resources from within *existing agency budgets*

- *No firm deadlines* to finalize & actually implement the public access plan
  - Key components don't need to 'kick in' until the plan is final

- Much *conditional phrasing*
  - Lots of '*shoulds*', not too many '*shalls*'

- Plenty of *caveats*
  - Lots of opportunities to alter, soften, or delay actions

Questions?

None? Good, let's get on with the discussion!

Open discussion

How might these requirements impact USGS? For example:

- Need for modification, expansion, or better integration of existing applications, systems, & services;
- Requirements for new applications or functionality;
- Altered workflows & management responsibilities;
- New or modified relationships with external partners (research, academic & commercial);
- Workforce needs including support training, education, realignment;
- Identification, protection & preservation of 'authoritative data';
- Imposition of new requirements (and burdens?) on scientists to speed use & release of their research & data;
- New policy requirements (e.g., data/publication release, RGE evaluation criteria);
- Budget impacts